Introduction to Long-Term User Experience Methods

Tiina Koponen, Jari Varsaluoma, Tanja Walsh
Seminar: How to Study Long-Term User Experience?
DELUX Project
1.6.2011

Unit of Human-Centered Technology
Department of Software Systems
Tampere University of Technology (TUT)
Contents

• "Longitudinal" paradigms in Human-Computer Interaction
• Methods for evaluating long-term UX
  • Diary Method
  • Experience Sampling Method
  • Day Reconstruction Method
  • Attrakdiff
  • Critical Incident Technique
  • UX curve
“Longitudinal” paradigms in HCI

Cross-sectional

Repeated sampling

Longitudinal

Retrospective
How do we determine if people enjoy what they do?

- **Study people’s choices**: they know what is good for them and reveal their preferences in their choice of activities (favored by economists).

- **Ask for self-reports of enjoyment** (favored by psychologists and social scientists):
  a) Global reports, for example, rating scales
  b) Assess people's momentary hedonic (pleasure-related) experience
  c) Collect retrospective reports of respondents’ feelings

Choosing methods for long-term user experience research

- Which UX factors are you interested in?
- Longitudinal or retrospective study?
- On-site or remote?
- Qualitative or quantitative data?
- Resources available?
  - time, budget, human resources
  - recruiting participants
  - analysing the data
  - utilising the results
Diary Method
A diary is a document created by an individual who maintains regular recordings about events in life, at the time that those events occur (Alaszewski 2006)

- Adapted to HCI from sociology and history
- Diaries allow to collect detailed user-defined research data
- Feedback diaries (events that interest researcher)
- Elicitation diaries (events that interest the user)
- Reduced retrospective reflection
Diary Method (2/5)

- Diaries are useful when:
  - Little is known about the usage patterns of a new technology
  - Technology is being used on the go
  - Research questions lead to data that cannot easily be observed or measured (such as feelings of frustration)
  - Recruitment of users and keeping them active diarists is challenging
Pros and Cons of Diary Method (3/5)

Pros
• Reduced impact of individual’s personal interpretation of what happened
• Qualitative & quantitative data
• Contextual & user-defined data
• Good for collecting data that changes over time e.g. moods
• Supports interviews
• Remote studies

Cons
• Harder to recruit participants than in surveys
• High drop out rate
• Problems with self reporting
• Analysis may take long
Tips (4/5)

- Good prizes enhance the answering activity
- Give clear instructions for participants of what they are supposed to record e.g. design the diary with some structure e.g. open questions
- Diaries work well with interviews
More information about Diaries (5/5)


Experience Sampling Method
Experience Sampling Method (1/5)

- Experience Sampling (ESM) is a method to collect information on people’s experiences in real time in natural settings immediately after the event during selected moments of the day.
- More reduced retrospective reflection than diaries.
- ESM collects information about both the context (physical and social) and content of daily life (activities, thoughts, feelings).
Experience Sampling Method (2/5)

- Responses for both open- and closed-ended questions at several random points during the day
- Minimizes the biasing effect of retrospection, but burdens participants as they are asked to interrupt their current activity
Experience Sampling Method (3/5)

As you were beeped...

Date ___ Time you were beeped ___ a.m./p.m. Time you responded ___ a.m./p.m.

Where were you?

What was the main thing you were doing?

What else were you doing at the same time?

What was on your mind?

Were you alone... Yes ___ No ___ or were you with...

(please check all that apply)

Your Spouse ___ Your Boss ___ Co-workers ___ Friend(s) ___ Girl/Boyfriend ___

Your Mother ___ Your Father ___ Teacher ___ Classmates/peers ___ Other(s) ___

Your child(ren) _______ (please indicate who)

Your sibling(s) _______ (please indicate who)

Think back on how you got into this activity...

Were you doing this main activity because you... (check all that apply)

(1) wanted to ___ (2) had to ___ (3) had nothing else to do ___

Indicate how you felt about the main activity. (please circle one number for each question)

<table>
<thead>
<tr>
<th>Not at all</th>
<th>A little</th>
<th>Somewhat</th>
<th>Very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you enjoy what you were doing?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Was this activity interesting?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>How well were you concentrating?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Were you living up to your own expectations?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Did you feel in control of the situation?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Did the situation allow you to be involved or to act?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Did you have the abilities to deal with the situation?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Was the activity important to you?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Were others expecting a lot from you?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Were you succeeding at what you were doing?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Did you wish you were doing something else?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Did you feel good about yourself?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

As you were being beeped, were you feeling...? (circle one number for each question)

<table>
<thead>
<tr>
<th>Not at all</th>
<th>A little</th>
<th>Somewhat</th>
<th>Very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheerful</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Lonely</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Nervous</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Cooperative</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Angry</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Responsible</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Frustrated</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Competitive</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Strained</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Caring</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Irritated</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Relaxed</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Stressed</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Proud</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Friendly</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Hardworking</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Productive</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Did you feel any physical pain or discomfort as you were beeped? (Please describe)

None ___ Slight ___ Bothersome ___ Severe ___

If you were talking with people, please answer the following 3 questions:

<table>
<thead>
<tr>
<th>Not at all</th>
<th>A little</th>
<th>Somewhat</th>
<th>Very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>Were you able to express your opinion?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Were others really listening to what you had to say?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Did you care about what others were saying?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

If you felt a strong emotion since the last report, what did you feel and why did you feel that way?

I felt ___ because ___

Please be specific

If anyone else you were with expressed a strong emotion to you since the last report, what did they feel and why?

(who) felt ___ because ___

Please be specific

Comments, etc.

* 500 Family Study (Schneider & Waite, 2005)
Pros and Cons of ESM (4/5)

Pros
- More accurate than diaries
- Rich data
  - Qualitative & quantitative
  - Contextual & user-defined
  - Emotions
- Field studies
- Quality of Everyday Life

Cons
- Demanding for participants
- High drop out rate
More Information about ESM (5/5)


Day Reconstruction Method
The Day Reconstruction Method (DRM) assesses how people spend their time and how they experience the various activities and settings of their lives (Kahneman et al., 2004)

In DRM participants...

1. Memorize the previous day by constructing a personal diary consisting of a sequence of episodes
2. Describe each episode by answering questions about the situation (when, what, where, with whom) and about the feelings that they experienced
### Retrospective:
“What happened today/yesterday?”

Memory-based evaluation of experiences may be biased, but these memories will be reported to others and guide the future behaviour of the individual. (Norman, 2009; Karapanos et al. 2010)

```
<table>
<thead>
<tr>
<th>Episode Name</th>
<th>Time it Began</th>
<th>Time it Ended</th>
<th>Notes to yourself: What happened?</th>
<th>What did you feel?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

First Morning Episode

Please look at your Diary and select the earliest episode you noted in the Morning.

When did this first episode begin and end (e.g., 7:30 am)? Please try to remember the times as precisely as you can.

This is episode number _____, which began at _______ and ended at _______.

What were you doing? (please check all that apply)

- commuting
- shopping
- doing housework
- eating
- praying/worshipping/meditating
- socializing
- napping/resting
- relaxing
- intimate relations
- other (please specify_________________)

Where were you?

- At home
- At work
- Somewhere else

Were you interacting with anyone (including on the phone, in a teleconference, etc)?

- no one → skip next question.

If you were interacting with someone (please check all that apply)

- spouse/significant other
- friends
- co-workers
- clients/customers/
- students/patients
- my children
- parents/relatives
- boss
- other people not listed

How did you feel during this episode?

Please rate each feeling on the scale given. A rating of 0 means that you did not experience that feeling at all. A rating of 6 means that this feeling was a very important part of the experience. Please circle the number between 0 and 6 that best describes how you felt.

<table>
<thead>
<tr>
<th>Feeling</th>
<th>Not at all</th>
<th>Very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impatient for it to end</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Happy</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Frustrated/annoyed</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Depressed/blue</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Competent/capable</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Hassled/pushed around</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Warm/friendly</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Angry/hostile</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Worried/anxious</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Enjoying myself</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Criticized/put down</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Tired</td>
<td>0</td>
<td>1 2 3 4 5 6</td>
</tr>
</tbody>
</table>

Many variations of the method are possible:

- Participants describe three *most impactful* experiences related to product use each day and write a story that describes in detail the situation, their feelings and their momentary perceptions of the product (Karapanos et al. 2009)

**Event Reconstruction Method (ERM)**

(Schwarz et al. 2009)

- Respondents memorise the “most recent” episode of the activity (e.g. product use) and then proceed as in DRM
Day Reconstruction Method (5/6)

**Pros**
- Many variations possible
- Remote studies
- Episodic format can reduce biases common for retrospective reports
- When compared to Experience Sampling, DRM…
  - imposes less respondent burden
  - does not disrupt normal activities

**Cons**
- Longitudinal studies (weeks/months) can be taxing for participants
- Most suitable for ready products
- Recruiting and keeping participants can be challenging during longitudinal studies
- Analysing use stories can be laborious
- ERM: How “recent” is recent enough for an accurate report?
Day Reconstruction Method (6/6)

- http://www.allaboutux.org/day-reconstruction-method

**DRM used for studying long-term UX:**

AttrakDiff is questionnaire measuring how attractive the product is in terms of usability and appearance.

AttrakDiff consists of 32 word pairs representing extreme opposites (e.g. good - bad) and 7-point scale between each word-pair.

Measures four UX attribute of the product:
- Practical quality (e.g. complicated – simple)
- Hedonic quality identity (unprofessional – professional)
- Hedonic quality stimulation (conventional – inventive)
- Attractiveness (rejecting – inviting)
## Evaluation of the product Flickr

With the help of the word-pairs please enter what you consider the most appropriate description for **Flickr**. Please click on your choice in every line!

<table>
<thead>
<tr>
<th>human</th>
<th>technical</th>
<th>isolating</th>
<th>connective</th>
<th>pleasant</th>
<th>unpleasant</th>
<th>inventive</th>
<th>conventional</th>
<th>simple</th>
<th>complicated</th>
</tr>
</thead>
<tbody>
<tr>
<td>practical</td>
<td>unprofessional</td>
<td>ugly</td>
<td>attractive</td>
<td>impractical</td>
<td>disagreeable</td>
<td>likeable</td>
<td>straightforward</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cumbersome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Pros**
- Provides comparative quantitative data
- Easy and quick to answer for respondents
- Quantitative data is quick to analyze
- Data can be collected remotely
- Can be used in evaluating any kinds of products

**Cons**
- Doesn’t provide qualitative data of the reasons why the product was attractive or unattractive and how it could be improved
- Doesn’t tell how important different UX elements are
- International studies: be careful when translating to different languages
- Assesses reflection on experiences, not actual experiences (all about UX)
More info:
http://www.attrakdiff.de/en/home/

AttrakDiff has been used, for example, in:
CIT- Critical Incident Technique (1/4)

• Method for collecting information of critical incidents
  • Critical incidents are either particularly **satisfactory** or especially **unsatisfactory** experiences.

• Focuses on important, detailed issues that are meaningful for the user/system/process, NOT on everyday habits or common practices.

• No structured questions - Users are simply asked to recall specific events and describe them in detail
  • CIT reflects the normal way users are thinking
  • Users are not forced into any specific framework
    > Provides pure user data

• Usually applied using interviews. Possible to use also questionnaires.
CIT- Critical Incident Technique (2/4)

• Why to ask just critical incidents/specific events?
  • Gives concrete information on details
  • Limited time-recourses: when there is no time to collect and/or analyze all incidents

• CIT reveals quality aspects relating to
  • Minimum product/service requirements
  • Value-enhancing aspects

• CIT provides information that guide consumer behavior *
  • Critical incidents are remembered for a long time - often for years
  • Critical incidents influence the customers company-directed behavior (praising, complaining, repeat purchasing) and third-party directed behavior (recommending, warning)

* (Hentchel 1992, Edvardsson 1988)
### CIT - Critical Incident Technique (3/4)

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Focuses on important issues; what is meaningful for the user</td>
<td>• Routine incidents may not be reported</td>
</tr>
<tr>
<td>• Useful for identifying rare events that might not be picked up by</td>
<td>• Critical incidents often rely on memory, incidents may be</td>
</tr>
<tr>
<td>other methods that focus on common or everyday events.</td>
<td>distorted or even forgotten if the incident is collected long after</td>
</tr>
<tr>
<td>• Provides information on details</td>
<td>the event</td>
</tr>
<tr>
<td>• Light for the user</td>
<td></td>
</tr>
</tbody>
</table>
CIT- Critical Incident Technique (4/4)

More info:


- http://www.usabilitybok.org/methods/p2052?section=basic-description
UX Curve

• UX Curve is retrospective method for evaluating long term UX – particularly **hedonic quality**.

• Starting point: to develop a lightweight method for industry to help understanding how hedonic aspects of **UX change over time**.

• User sketches a curve of how his/her relationship with the product has changed over time.
  • User Curve starts from the moment of purchase and continues till present.
  • Besides drawing a curve, user marks peaks and lows to the curve and explains what happened in these points: What caused the change, why the relationship changed e.g. decreased heavily
UX Curve

- Curves can be drawn of different dimensions of UX:
  - Attractiveness
  - Ease of use
  - Utility
  - Usage volume
  - "General curve"

- Analysis
  - Content analysis
  - Curve shape: Ascending/Descending curves

Please, recall the moment when you started to use the product and draw a curve describing how your relationship has changed from the first time till today.

Your relationship towards the product

Please, describe the reasons for the changes in your relationship:

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
Pros
• Data for a long period of time can be collected in one go
• No need to commit users for weeks/months/years
• Provides qualitative data on what is meaningful and important for the users
• Provides information of hedonic (pleasurable) UX aspects

Cons
• Relies on memory, details may be distorted or even forgotten if the details collected long afterwards
• New method:
  • Not much experience when the data should be collected
  • Not much experience how UX curve works in different contexts
  • Analysis somewhat challenging (we are working on these issues)
<table>
<thead>
<tr>
<th>Method</th>
<th>What is this about?</th>
<th>Where to use?</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIARY</td>
<td>Reporting events of daily life for days or weeks</td>
<td>Supports interviews, understanding context</td>
</tr>
<tr>
<td>EXPERIENCE SAMPLING</td>
<td>More specified data of certain events, days or weeks</td>
<td>Moods, quality of daily life, specific usage experiences</td>
</tr>
<tr>
<td>DAY RECONSTRUCTION</td>
<td>Summary of daily events, weeks</td>
<td>Well-being studies, emotions for longer periods</td>
</tr>
<tr>
<td>ATTRAKDIFF</td>
<td>Summative quantitative evaluation of product’s attractiveness</td>
<td>Any cases, together with other methods</td>
</tr>
<tr>
<td>CRITICAL INCIDENT TECHNIQUE</td>
<td>Particularly satisfactory or unsatisfactory experiences during longer time periods</td>
<td>Quality service research, peak experiences</td>
</tr>
<tr>
<td>UX CURVE</td>
<td>How UX changes over time and what causes changes over months</td>
<td>Predicting and supporting long-term UX and customer loyalty</td>
</tr>
</tbody>
</table>
Questions? Comments?

Thank you!

Contact us: delux@cs.tut.fi